DOCUMENT RESUME

ED 108 439 EC 073 163

AUTHOR	Corman, Louise; Budoff, Milton
TITLE	The Picture Word Game: A Nonverbal Test of the Ability to Use Language-Related Symbols. Vol. 4, No. 77.
INSTITUTION	Research Inst. for Educational Problems, Cambridge, Mass.
PUB DATE	74
NOTE	33p.
EDRS PRICE	MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS	Disadvantaged Youth; *Economically Disadvantaged; Elementary Education; Exceptional Child Research; *Language Handicaps; *Language Tests; *Nonverbal Tests
IDENTIFIERS	Picture Word Game

ABSTRACT

The Picture Word Game, a nonverbal test of the ability to use language-related symbols, was administered to 90 low-income urban students in grades 1 through 5. Ss were trained in one class period and then tested in class size groups within 2 days. In addition, previous scores on the Stanford Paragraph Meaning and Vocabulary Subtests were obtained from school records. Results indicated that the training-based procedure had advantages for assessing the language skills of low-income students who often have verbal deficits which inhibit them from displaying maximal competence on traditional tests. (Appendixes contain sample copies of the student's training booklet and of the final version of the test.) (LH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED108439

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE PICTURE WORD GAME: A NONVERBAL TEST OF THE ABILITY TO USE LANGUAGE-RELATED SYMBOLS[1]

Louise Corman and Milton Budoff
Research Institute for Educational Problems

Budoff and his associates have developed a training-based procedure for assessing reasoning ability on nonverbal tasks (Budoff, 1970). In this assessment strategy training assumes a critical role, particularly for the child from a poor and/or nonwhite background who may learn different cognitive strategies in expressive formats other than those presumed to be available by traditional tests. The training helps the child to narrow the cognitive gap between his previously learned problem-solving strategies and those implicit to the problems he must ordinarily solve on the middle-class-biased tests he encounters. Inclusion of training in the assessment procedure also minimizes the artificiality of the test situation. Repeated contacts with the materials in a context of support and teaching allow the school-failing child to develop a sense that he can be competent. Without this competence boost, he tends not to perform at his best, implicitly expecting failure (Zigler, 1966)

The essence of this assessment strategy, then, is to impose some control on the potentially negative effects on the child's test performance of prior life experiences. Two types of effects can be considered: those which are due to problem solving experiences which differ from the abstracting verbal conceptual types of skills expected of school children, and the negative effects of failing to perform well on the tests the child has taken during his school years. Test scores after training should reflect the child's ability under optimized conditions in which he is familiar with the task and its demands, has had success in solving problems similar to those on the test, and has had the opportunity to learn and apply relevant strategies.

The goal of this study was to apply the rationale from training based assessment to tasks that are language related, since previous efforts had focused on nonverbal reasoning tasks. The specific aim was to develop a procedure that would permit the low SES, low school achieving child to show his/her proficiency in utilizing language related symbols, following a systematic learning experience. In this task, the child learns a new symbol system for common words that he understands, and must apply this understanding in generating or decoding sentences. A critical feature of this language measure is that it not make excessive demands on the previously developed verbal skills of students who may have experienced difficulty with verbal tasks.

The Picture Word Game (PWG) was conceived as a modification of the Semantic Test of Intelligence (STI) constructed by Philip Rulon. The STI is a language-related measure in which the examinee learns to associate a geometric symbol with a pictured object or action (man, woman, runs, pushes). The student must "read" the sequence of symbols and choose the appropriate picture from a multiple choice array. The task is an analogue to reading the words in a sentence and indicating the picture that best represents the meaning of the sentence. As such, it was thought that the task might represent a means by which one could examine verbally related competencies of children in a training-based format, without the necessity of verbal expressive materials or reading words, skills in which these children are often deficient.

The STI is administered as a timed test and consists of 217 items, including 109 items with one symbol, 49 with two symbols, 36 with three symbols, and 23 with four symbols. The symbols for each noun and verb are introduced as single symbols, and defined by accompanying pictures in a multiple choice format on tuition pages which are not scored. Tuition pages are used to introduce 2-, 3-, and 4-symbol "sentences." Instructions are pantomimed. The symbols and the pictures which define their meaning are presented on each double page so that memory for the meaning of the symbol is not a factor influencing performance.

The STI was developed as a measure of military trainability for illiterate recruits who had failed the literacy requirement for entry into the Marine Corps. Validity of the test as a language measure was evidenced by Rulon and Schweiker's (1953) finding that recruits who did well on the STI also successfully completed a literacy course to meet the eligibility requirements of the Marine Corps.

In a previous study (Gimon, Budoff, & Corman, 1974), the investigators administered the STI to 76 Spanish-speaking children from low income families, who ranged in age from 6 to 13 years. Results indicated that a ceiling effect occurred with children over 8 years old. The mean number of one-symbol items passed by over 50% of the total sample was 92 of 109 (84%), and all but seven of the 49 two-symbol items were passed by more than half the children. These findings revealed that the majority of items on the test were too easy for the children in this age range. This investigation also revealed that STI scores were significantly correlated with vocabulary scores on both the Spanish and English versions of the WISC. The investigators concluded from these results that the tasks used in the STI were related to verbal ability, but that the difficulty level of the test would have to be broadened in order for it to be used with normal children spanning a wider age range.

The Picture Word Game was constructed to provide a minimally verbal measure of language ability of children in the first through fifth grades. Development of the test was conducted in three phases: i) construction of the original test and accompanying training procedure, ii) test development, i.e., revision of the original test, and iii) evaluation of psychometric characteristics of the final test.

Construction of the Original Test

The following principles represent departures from the STI format and were used with the Picture Word Game in an effort to broaden the difficulty level of the STI:

1. The Picture Word Game is a power test, administered as an untimed measure. Eliminating the effect of speed in performance was considered desirable to reduce spurious inflation of variability and consequently reliability.

2. The vocabulary in the test, i.e., the number of symbols, was increased from nine "words" in the STI to 16 in the Picture Word Game.

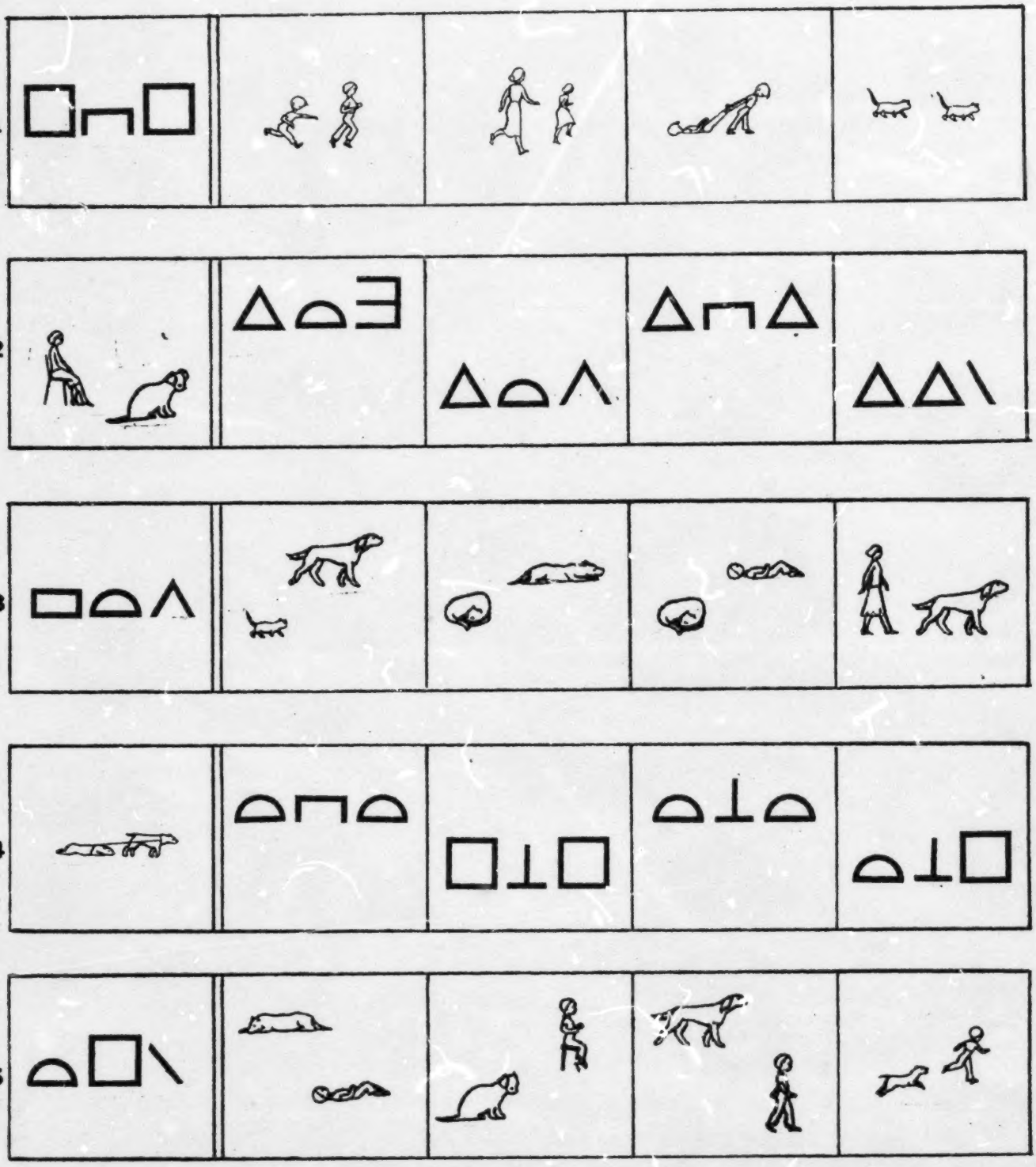
3. The concept of symbols representing numbers was included, in addition to noun and verb representation.

4. Whereas all items in the STI require translation from symbol to picture, approximately half the items in the Picture Word Game require translation from picture to symbol. The investigators believed that translation from picture to symbol would be a more difficult task, especially when three or more symbols were to be read. An internal measure of the effectiveness of the teaching sequence would be the extent to which the child can utilize the symbol in both directions--to decode the symbols into their pictorial equivalents, and to generate sentences in the "new" language that would explain the picture. (See Figure 1).

5. While the STI contained 217 items, including a large number of very easy one and two symbol items, the original version of the Picture Word Game contained 60 items with a much smaller proportion of one and two symbol items. The intent was to reduce the number of items on the final form even further so that the Picture Word Game could be administered in one class period.

6. While the most difficult items on the STI require reading a "sentence" of four symbols, the Picture Word Game included items of five symbols as well. It was not possible to construct a meaningful sentence of more than five symbols within this format, when these symbols represented only nouns, verbs, and numbers.

FIGURE 1



The original form of the Picture Word Game contained 60 items, with odd-numbered items requiring translation from symbol to picture and even-numbered items requiring translation from picture to symbol. The test included six one-symbol, 12 two-symbol, 18 three-symbol, 18 four-symbol, and 6 five-symbol items. Vocabulary comprised 16 symbols: five nouns (cat, dog, woman, boy, horse), five intransitive verbs (sit, lie down, walk, stand, run), three transitive verbs (pull, chase, carry), and three numbers (two, three, and four).

A training procedure was developed such that training could be entirely concluded before administration of the test, rather than embedding training within the test as in the STI. Oral instructions were used in the training, instead of relying solely on pantomimed instructions. The latter procedure used in the STI was considered artificial by the investigators on the basis of a previous administration of the STI (Gimon, Budoff, & Corman, 1974).

A training booklet was devised which contained 10 of the 16 symbols used in the vocabulary of the test. A page with all the symbols in the vocabulary was attached to the back of the training booklet, so that it could be easily detached and referred to by the student during the training session. The six symbols contained in the test but excluded from training were one noun, one intransitive verb, one transitive verb, and the three numbers. The training booklet contained items similar to those in the test, and slides were made of each training item. The trainer simultaneously displayed the slides while explaninig the principles of the tasks and techniques for solving them to students using the training booklet.

Test Revision

Subjects and Procedure

Pilot testing was conducted in June of 1973, with a sample of 205 students from a low income school district in an urban community in Massachusetts. The subjects constituted ten first through fifth grade classrooms, two classes per grade, and were evenly divided by sex. The mean age of the sample was 9 years, 2 months, with a standard deviation of 1 year, 7 months.

All students were trained in one class period and tested in class size groups in a second class period three days following training. A student's score was calculated as the percent of items he answered correctly of the total number of test items.

Results

The test showed a high degree of internal consistency reflected in a KR20 reliability coefficient of .95. The mean

percent of correct responses for the total sample was 75.5 (SD = 20.8). The mean percent correct for students in grades one to five, respectively, were: 62.7 (±21.3), 63.0 (±21.2), 86.9 (±11.8), 79.5 (±11.8), and 89.1 (±11.8). These figures indicated that, despite efforts to make this test more difficult than the STI, a ceiling effect was found for students at or beyond the third grade. Mean scores by grade level reflected a dichotomy between first and second graders on the one hand and third through fifth graders.

Selection of Items for the Final Test

Difficulty levels and discrimination indices were obtained for each item with the total sample and subjects in each grade. Items were selected for the final test which met three criteria: i) the difficulty level of test items for the total sample was evenly distributed throughout the test, i.e., approximately equal numbers of easy and moderate-to-hard items were retained, ii) to the extent it was possible, each item selected showed a gradual increase in difficulty from grades one to five, and iii) the discrimination of each item was not less than .25 for the total sample. These procedures were used in an attempt to produce test scores that would reflect a developmental trend and reduce the likelihood of a ceiling effect on the final test. The revised test consisted of 37 items, including one one-symbol, 7 two-symbol, 9 three-symbol, 9 four-symbol, and 11 five-symbol items.

Other revisions based on pilot testing consisted of improvements in pictures used in specific items. An effort was made in both the test and training booklets to sharpen pictures which the children had difficulty interpreting. Appendix A contains the final training instructions and student's training booklet. The final version of the test is presented in Appendix B.

Characteristics of the Final Test

Subjects

The sample for this phase of the study, which was conducted in January, 1974, consisted of 90 students from a low income area in Massachusetts. Subjects were evenly divided into five first through fifth grade classrooms, one class per grade, and attended an elementary school in the same urban community as the school from which the pilot sample was drawn. Subjects were evenly divided by sex. The mean age of the sample was 8 years, 10 months (SD = 1 year, 9 months).

Procedure

All students were trained in one class period and tested in class size groups within two days following training. In addition, all students' scores on the Stanford Paragraph

Meaning and Vocabulary Subtests were obtained. The Primary II level of this test had been administered to third through fifth graders by the school when these students were in the third grade, and scores were obtained from their school records. First and second graders were group administered the Primary I level after Picture Word testing had been completed.

Results

Discrimination and difficulty levels of final test items are presented in Table 1. Discrimination indices of almost all items were high, and this fact was reflected in the high KR20 reliability coefficient of .93. Mean item difficulties revealed that several procedures used to modify the STI were successful in increasing the difficulty level of the Picture Word Game when the total sample is considered: a) items requiring translation from symbol to picture were significantly harder than items which required translation from picture to symbol ($F = 63.64$, 1/85 df, $p < .001$); b) the difficulty of items increased linearly as the number of symbols increased from two to five ($F = 70.27$, 3/255 df, $p < .01$), and c) the two most difficult sets of items on the test were items employing the concept of numbers which had not been taught, and all of which had three or more symbols (mean = 59.3%) and the five symbol items (mean = 53.4%). The STI does not include either of these item types.

Despite the effectiveness of these procedures, however, mean total scores (in terms of percent correct) for each grade again revealed a ceiling effect for children beyond the third grade whose average score exceeded 30% correct (Table 2). Total score means reflected the same dichotomy between grades 1 to 2 and 3 through 5 which was found on the pilot test, despite item selection procedures which retained items with the most marked linear developmental trend. The table indicates that the gap between second and third graders widened as the number of symbols increased from three to four. Items involving number concepts and those requiring translation from picture to symbol also differentiated first and second graders from the rest of the sample.

Mean stanines on the Stanford Comprehension and Vocabulary subtests were relatively comparable among the five grades and revealed that these children's reading skills were in the moderate to low range in relation to national norms. The mean stanine for the total sample was 4.3 (±1.8) on comprehension and 4.1 (±1.4) on vocabulary. Scores on the Picture Word game were correlated with scores on these two subtests to provide evidence of validity of the Picture Word Game as a language measure. Coefficients of .37 and .34 were obtained with comprehension and vocabulary, respectively. While these coefficients

TABLE 1

Discrimination and Difficulty Levels of
Final Picture Word Game Items

Item	Number of Symbols	Stimulus Mode[a]	Number Concept?	Discrimination	Difficulty
1	1	P		.04	.96
2	2	S		.29	.90
3	2	P		.45	.81
4	2	S		.32	.92
5	2	S		.27	.92
6	2	P		.43	.84
7	2	S		.50	.92
8	2	S		.44	.90
9	3	P		.44	.79
10	3	S		.46	.79
11	3	P		.49	.62
12	3	S		.57	.88
13	3	P		.73	.80
14	3	S	yes	.56	.61
15	4	P	yes	.44	.50
16	3	S	yes	.65	.58
17	3	P	yes	.47	.63
18	3	S		.62	.87
19	4	P		.47	.77
20	4	S	yes	.51	.65
21	4	P	yes	.46	.14
22	4	S	yes	.72	.77
23	4	P		.39	.63
24	4	S	yes	.74	.77
25	4	P	yes	.59	.68
26	4	P		.62	.63
27	5	P	yes	.67	.58
28	5	S		.59	.67
29	5	P		.65	.70
30	5	S	yes	.74	.61
31	5	P	yes	.62	.41
32	5	S	yes	.67	.64
33	5	P		.62	.38
34	5	S	yes	.75	.58
35	5	P		.58	.43
36	5	S	yes	.57	.43
37	5	P		.58	.44

[a]Items marked "P" require translation from picture to symbol; items marked "S" require translation from symbol to picture.

TABLE 2

Mean Percent Correct by Grade on the Picture Word Game

	Total Test		Number of Symbols								Stimulus Mode			
			Two[a]		Three		Four		Five		Picture to Symbol		Symbol Pict	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	
Total Sample	68.8	23.4	89.7	16.2	73.0	26.7	64.9	29.2	53.4	35.0	63.5	24.5	74.5	2
Grade 1	45.0	21.1	74.4	23.2	44.4	30.2	46.1	32.4	23.2	24.5	40.3	20.5	50.0	2
Grade 2	55.2	19.5	92.3	11.9	65.1	23.1	48.7	33.2	25.5	26.9	47.4	21.1	63.5	1
Grade 3	80.4	11.5	93.3	9.0	85.2	11.2	71.9	9.8	73.9	23.0	73.7	13.0	87.4	1
Grade 4	84.5	10.0	95.3	10.7	86.1	8.3	84.0	13.6	75.6	21.7	81.9	9.8	87.2	1
Grade 5	87.7	8.5	95.8	5.9	92.0	12.2	82.1	14.4	82.8	13.2	83.0	11.2	92.6	
Number of Items	37		8		9		9		11		19		18	

[a]Item 1 with one symbol was included in this group.

TABLE 2

Mean Percent Correct by Grade on the Picture Word Game

st D	Number of Symbols								Stimulus Mode						N
	Two[a]		Three		Four		Five		Picture to Symbol		Symbol to Picture		Number Concept		
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	
.4	89.7	16.2	73.0	26.7	64.9	29.2	53.4	35.0	63.5	24.5	74.5	24.1	59.3	30.5	90
.1	74.4	23.2	44.4	30.2	46.1	32.4	23.2	24.5	40.3	20.5	50.0	24.7	33.0	23.6	20
.5	92.3	11.9	65.1	23.1	48.7	33.2	25.5	26.9	47.4	21.1	63.5	19.8	37.1	25.9	21
.5	93.3	9.0	85.2	11.2	71.9	9.8	73.9	23.0	73.7	13.0	87.4	14.0	75.6	20.0	15
.0	95.3	10.7	86.1	8.3	84.0	13.6	75.6	21.7	81.9	9.3	87.2	11.7	77.9	16.6	16
.5	95.8	5.9	92.0	12.2	82.1	14.4	82.8	13.2	83.3	11.2	92.6	7.6	84.1	13.9	18
	8		9		9		11		19		18		15		

[a]ne symbol was included in this group.

are low in terms of concurrent validity, they are similar in magnitude to the validity coefficient of .385 obtained by Rulon and Schweiker (1953) with the STI.

Discussion

A training-based procedure clearly has advantages for assessing language skills of low SES children. These children typically have a verbal deficit which inhibits them from displaying their maximal competence on traditional verbal tests. Training provides a supportive context which can permit them to demonstrate optimum performance.

The basic principles employed in the STI would appear to lend themselves to adaptation for use as a language measure with children of a broad age range. Although certain test construction procedures were effective in increasing the difficulty of the STI, attempts to broaden the difficulty level of the test as a whole were largely unsuccessful. Like the STI, the Picture Word Game was shown to be too easy for children beyond the third grade, despite the fact that comprehension and vocabularly skills of this sample were below average on national norms.

In its present form the Picture Word Game appears to be most useful with second graders. It could be argued that the training session was so beneficial in maximizing test performance that training was partially responsible for the ceiling that resulted in the test. Some might suggest that training only items which require translation from symbol to picture might lower the test ceiling and at the same time permit examination of transfer of training to the more difficult picture to symbol items on the test. It is doubtful, however, that this procedure would effectively reduce the test ceiling; despite the fact that items involving number concepts or five symbols were excluded from the training, children beyond the third grade had little difficulty in translating such items on the test.

Results of this study indicated that the Picture Word Game measures a unitary ability which is related to language skills. It is likely that facility in translation required by this test is mastered by children of normal intelligence after the third grade and that, within the STI format, quantitative modifications are not sufficient to tap the increasingly complex language skills learned by children in the intermediate grades. Use of the Picture Word Game in further research with educable mental retarded children might prove to be quite fruitful; it is unlikely that the ceiling effect would be evident with these children for whom a minimally verbal, training-based measure is clearly appropriate and could serve a critical need.

References

Budoff, M. Learning potential hypothesis: Its effects on teacher expectancies for disadvantaged students. Paper presented at the Conference on Teacher Expectancies, Georgia State University, Atlanta, Georgia, October, 1970.

Gimon, A., Budoff, M., & Corman, L. Applicability of learning potential measurement with Spanish-speaking youth as an alternative to IQ. Final Report, 1974. Grant OEG-1-72-0020 (509), Office of Education.

Rulon, P., & Schweiker, R. Validation of a nonverbal test of military trainability. Final Report, June, 1953, Department of the Army.

Zigler, E. Mental retardation: Current issues and approaches. In L. W. Hoffman & M. L. Hoffman (Eds.), Review of child development research, Vol. 2. New York: Russell Sage Foundation, 1966.

Footnotes

[1]This research was supported by grant NE-G-00-3-0016 from the National Institute of Education, Department of Health, Education, and Welfare.

NAME _____ DATE _____

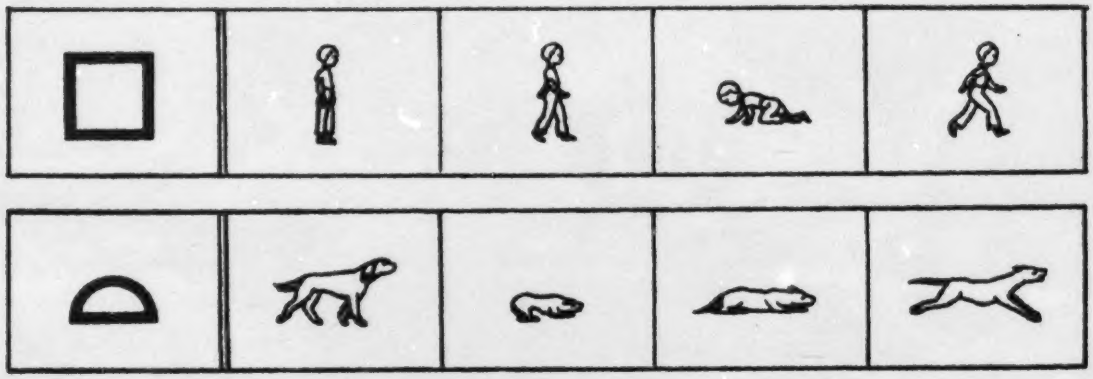
GRADE _____ TEACHER _____

SCHOOL _____

BIRTHDATE (MONTH, DAY, AND YEAR YOU WERE BORN) _____

SEX (CHECK ONE) _____ BOY _____ GIRL

THE PICTURE-WORD GAME



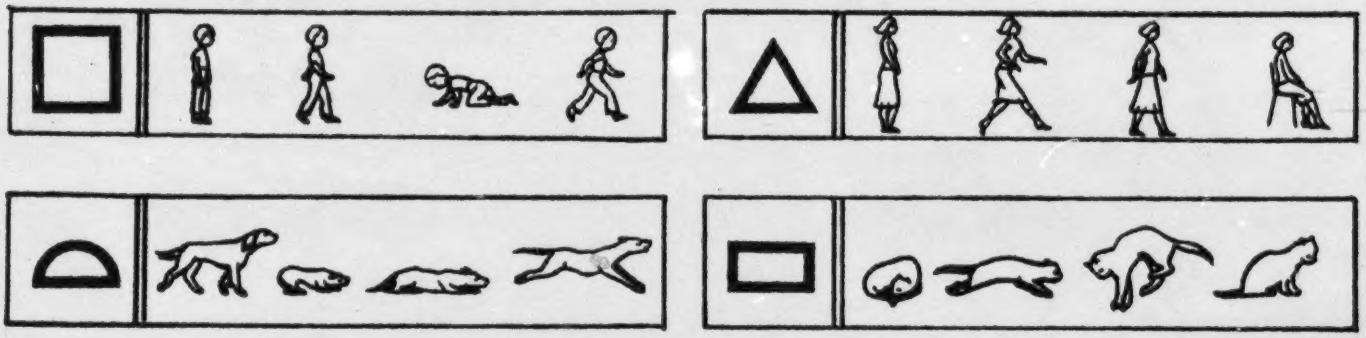
1

2

3

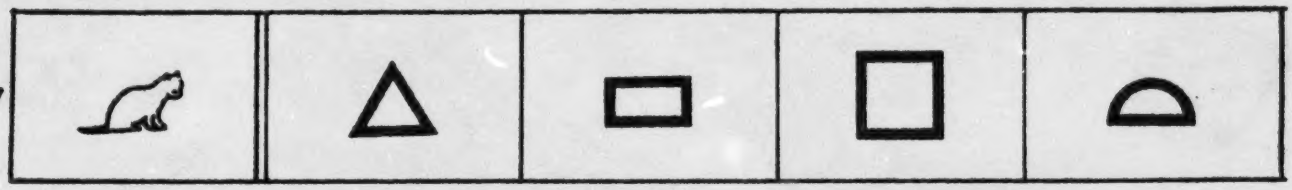
4

5



6

7



8

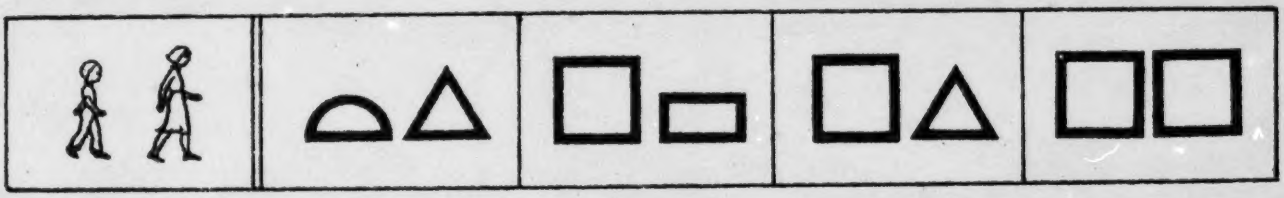
9

10

17

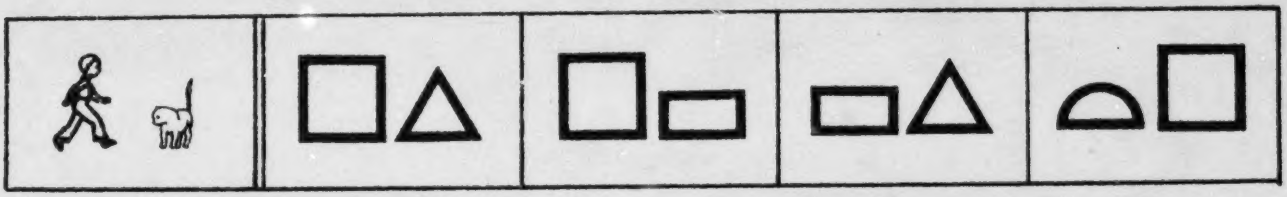
11

12

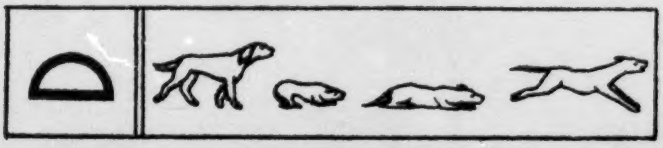


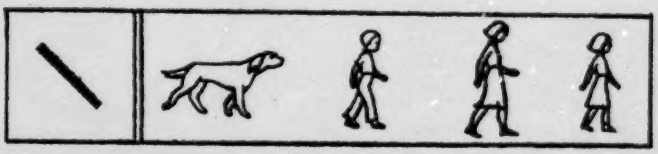
13

14



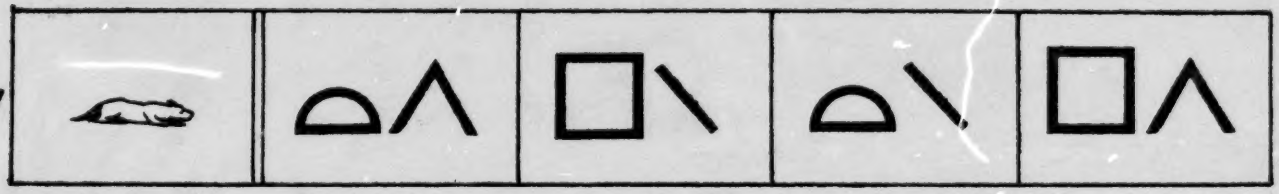
15





16

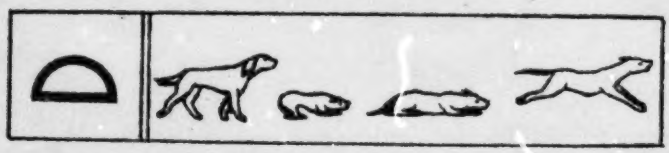
17

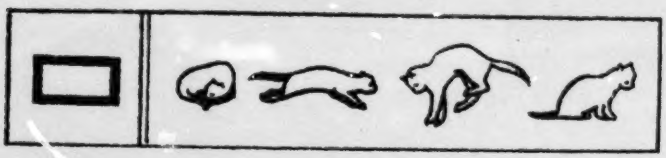


18

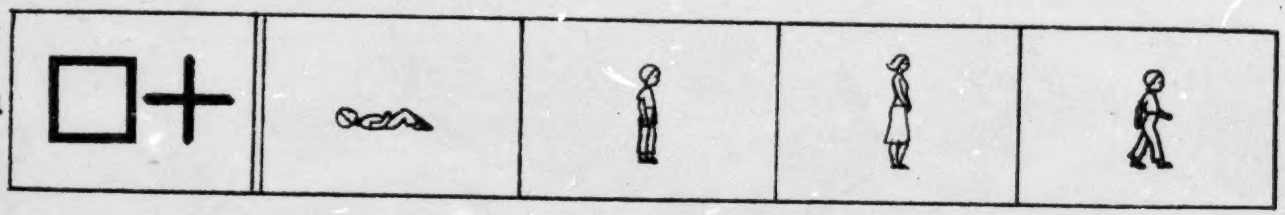
19

20

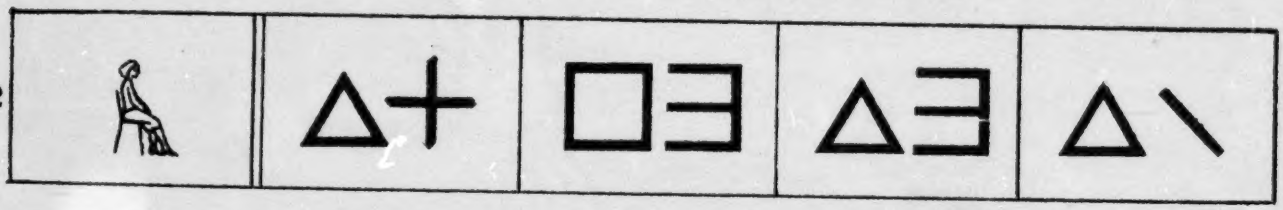




21

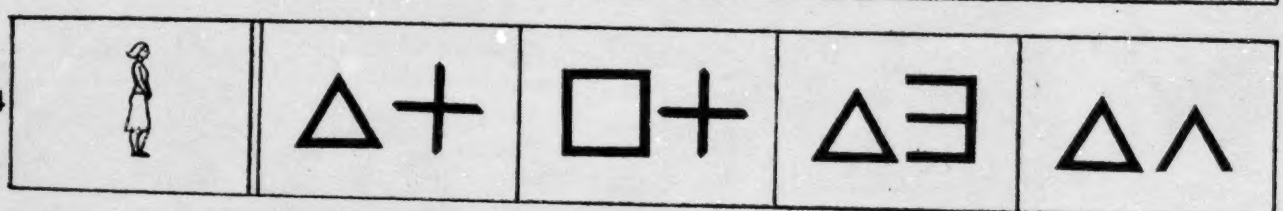


22



23

24



25

20

26

27

28

29

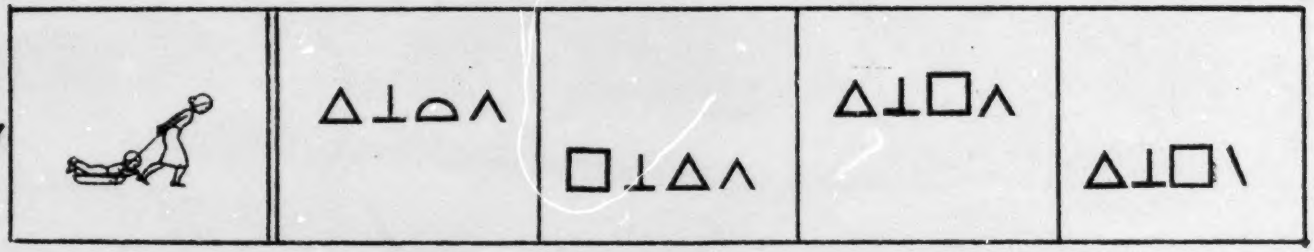
30

21

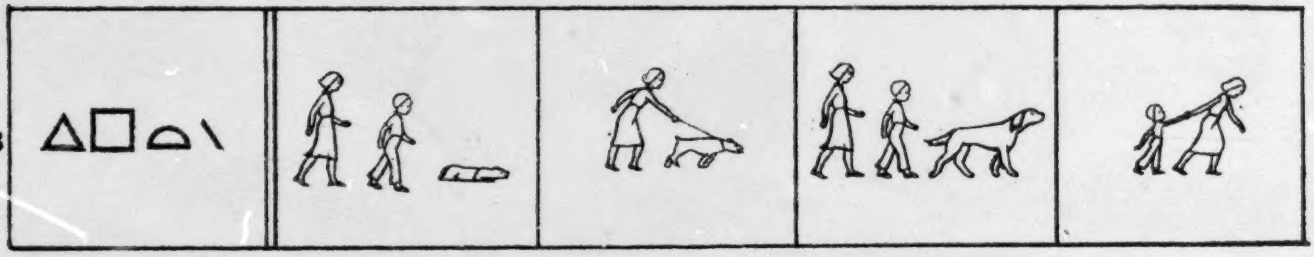
31
32
33
34
35

36

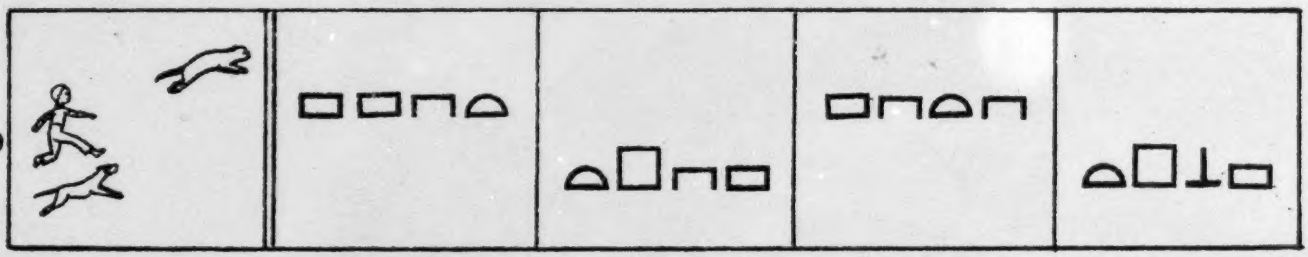
37



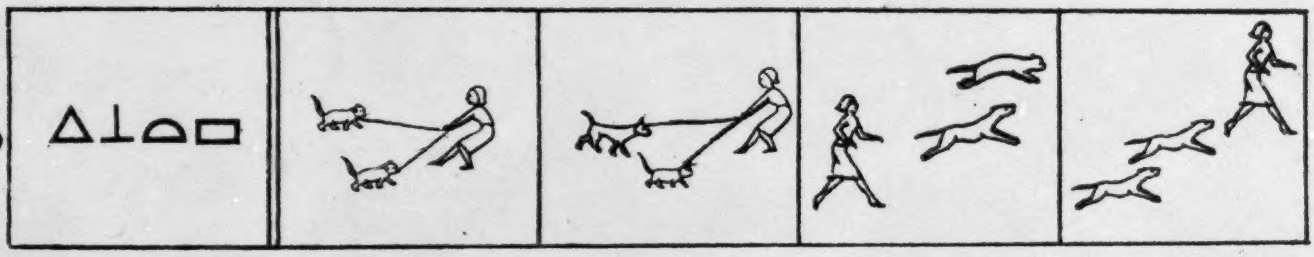
38

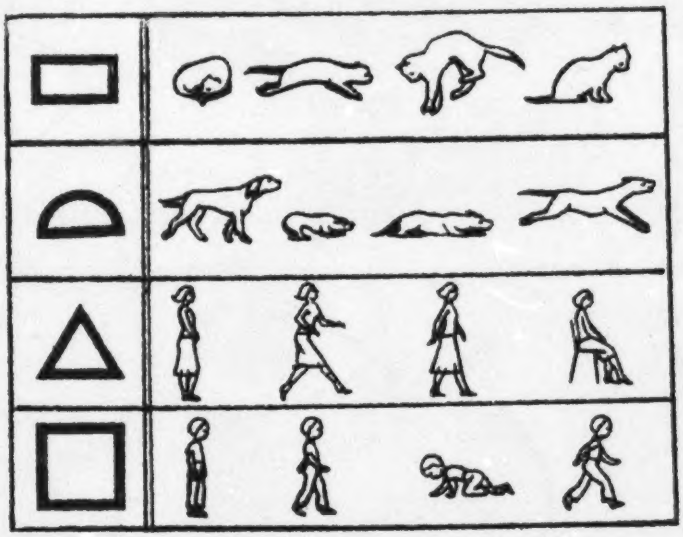


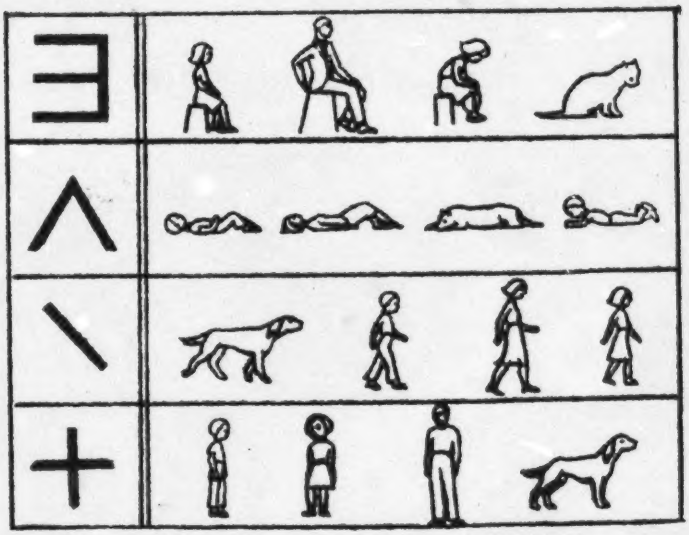
39

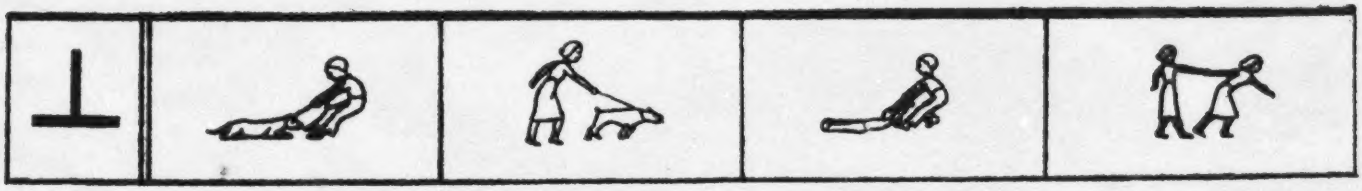


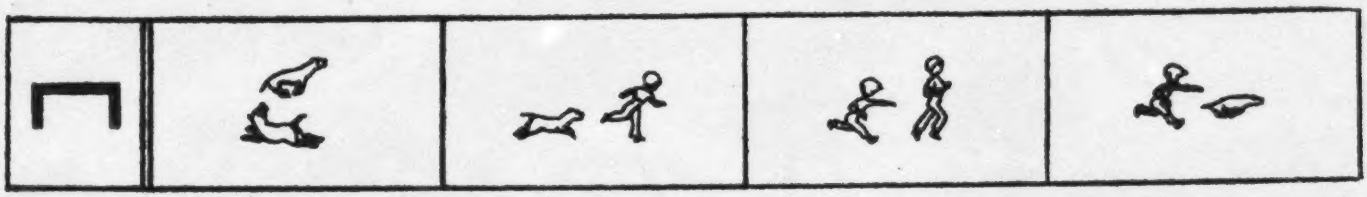
40











NAME _____ DATE _____

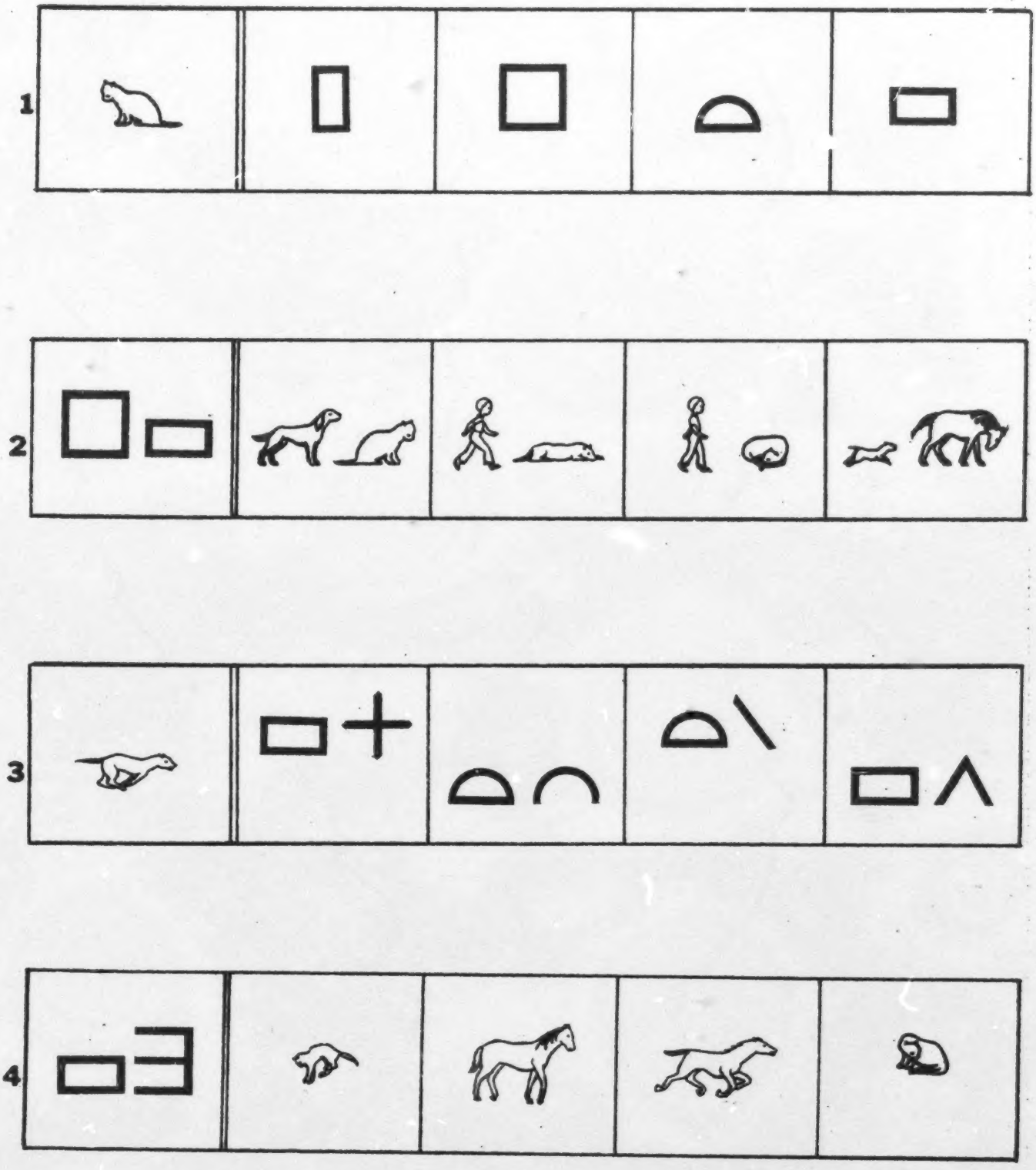
GRADE _____ TEACHER _____

SCHOOL _____

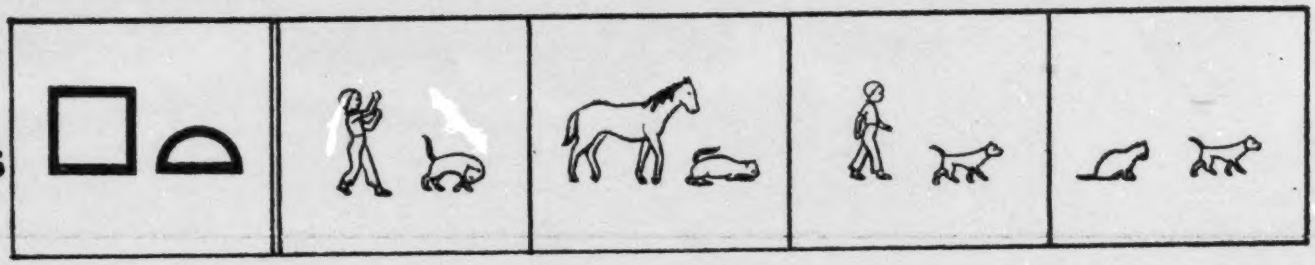
BIRTHDATE (MONTH, DAY, AND YEAR YOU WERE BORN) _____

SEX (CHECK ONE) _____ BOY _____ GIRL

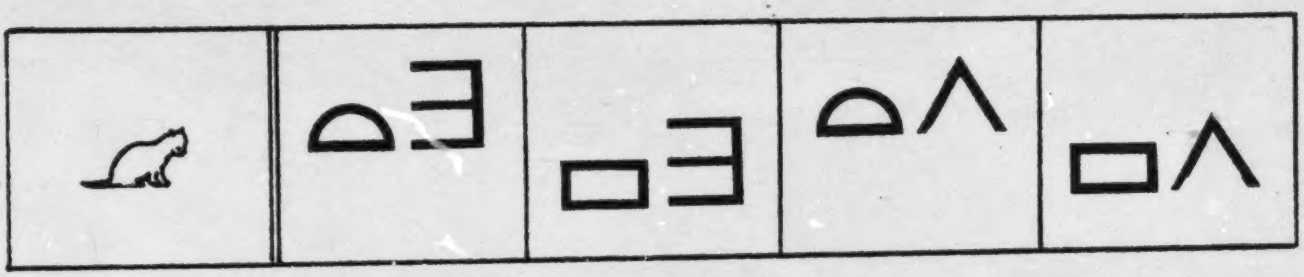
THE PICTURE-WORD GAME


1
2
3
4

5



6



7

8

27

9

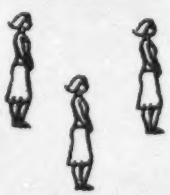
10

11

12

13

14



28

15

16

17

18

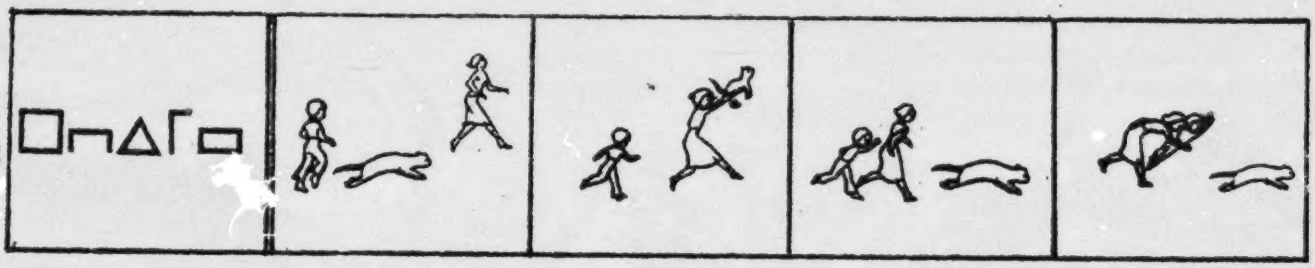
19

20

21					
22					
23					
24					
25					
26					

27

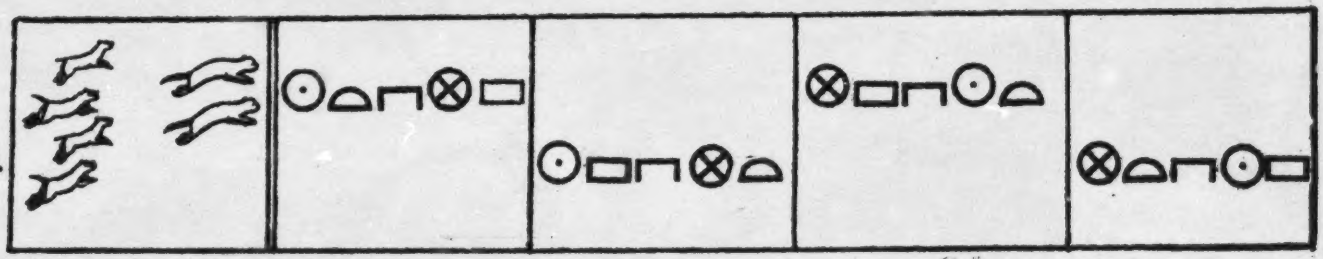
28



29

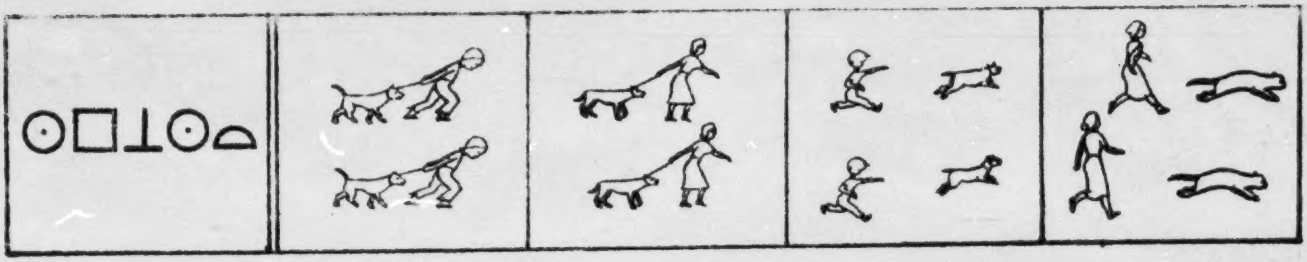
30

31

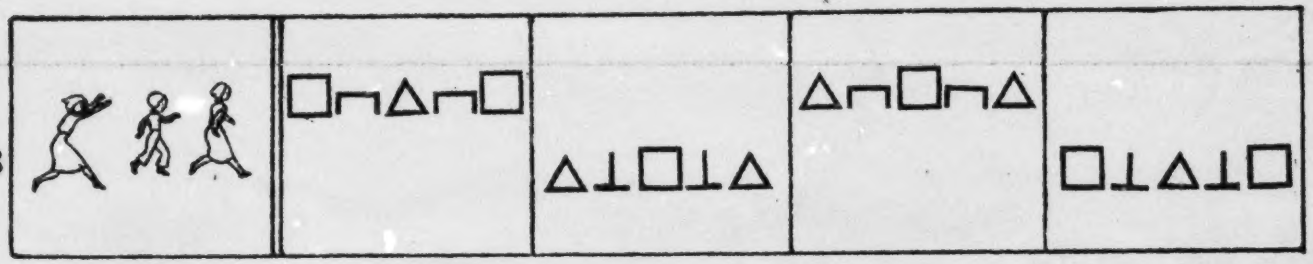


31

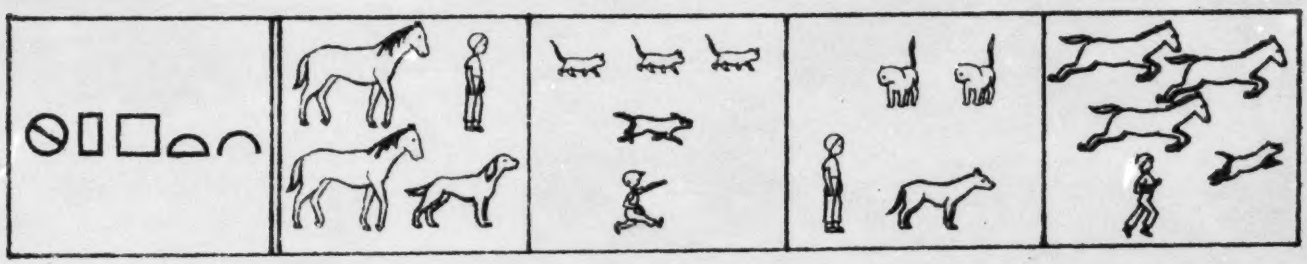
32



33



34



35

36

37